Multi-attribute Search and Informative Advertising

Yunfei (Jesse) Yao*

(University of California, Berkeley)

May, 2023

* I thank J. Miguel Villas-Boas, Yuichiro Kamada, Zihao Zhou, and seminar participants at the Berkeley/Columbia/Duke/MIT/Northwestern IO Theory Conference and Berkeley Econ Theory Lunch for constructive comments. All errors are my own. E-mail address: jesseyao@berkeley.edu.

Abstract

When considering whether or not to buy a product, the consumer can often evaluate different attributes of it. Sometimes, the consumer chooses which attribute to search for because of exogenous reasons (e.g., one attribute is more important than others). However, the consumer often is unclear about which attribute is more important ex-ante. Assuming that a product has two symmetric attributes, we study the optimal search strategy of the consumer by endogenizing the optimal attribute to search, when to keep searching for information, and when to stop searching and make a decision. We characterize the search region by a set of ordinary differential equations for moderate beliefs and by a system of equations for extreme beliefs. We find that it is always optimal for the consumer to search for the attribute about which she has the higher uncertainty due to the faster speed of learning. The consumer only searches for the more uncertain attribute if she holds a strong prior belief about one of the attributes, and may search for both attributes otherwise. We then show how firms can influence consumers' search behavior and increase profits through informative advertising. The firm does not advertise if the consumer's prior beliefs about both attributes are extreme. Otherwise, the firm advertises the better attribute if the consumer is optimistic enough about the worse attribute, and advertises the worse attribute if the consumer is less optimistic about it.

1 Introduction

When considering whether or not to buy a product, the consumer can often evaluate different attributes of it. An incoming college student finding a laptop can learn about the operating system, weight, exterior design, warranty, and many other attributes before making the final decision. Learning costs both time and effort, while consumers often have limited attention. So, she needs to decide which attribute to learn (first). Sometimes, she makes the decision based on exogenous reasons, such as an attribute is more prominent (Bordolo et al. 2013, Zhu and Dukes 2017), or her options in an attribute generate a greater range of consumption utility (Kőszegi and Szeidl 2013).

Consider a consumer deciding whether to buy a used car. She can gather information about many different attributes. Figure 1 summarizes some factors of the used car value. For example, the consumer can check details about the car's add-on packages by some review articles. She can also purchase a car report to find out the car's accident history. Both options help the consumer learn more about the car and improve the decision. However, it takes time and effort to search for such information. The consumer needs to decide to which attribute to pay attention. If one attribute is much more important than the other, she will search for information about the most important attribute. However, the consumer may not know whether the add-on features or the condition of the car matter more to her. What attribute should she search for if she is unclear about which one is more important ex-ante?

Even if the consumer decides which attribute to search for, she will not learn everything about it immediately. Instead, she gradually gathers information about the attribute. For instance, Even if the consumer spends half an hour searching for information about the car's safety features and finds out that the car has airbags in each of the seats, she still does not know everything about the car's safety. She can continue searching for information about whether the car has an automatic braking system. But, the consumer may not want to stick to one attribute. The relative importance of attributes may change as she learns more. After obtaining enough positive information about the car's safety, she may find it a better use of her time to switch to other attributes. She may feel confident that the car is safe but uncertain whether she will enjoy driving in it. At some point, the consumer may switch to learning more about the car's design. When will the consumer switch to search for another attribute because the relative importance of attributes changes as she gathers more information?

This paper considers a consumer deciding whether to purchase a good or not. The good has two attributes, whose values are independent. The payoff of purchasing the good is the total value of the attributes net of the price. The consumer does not know the value of

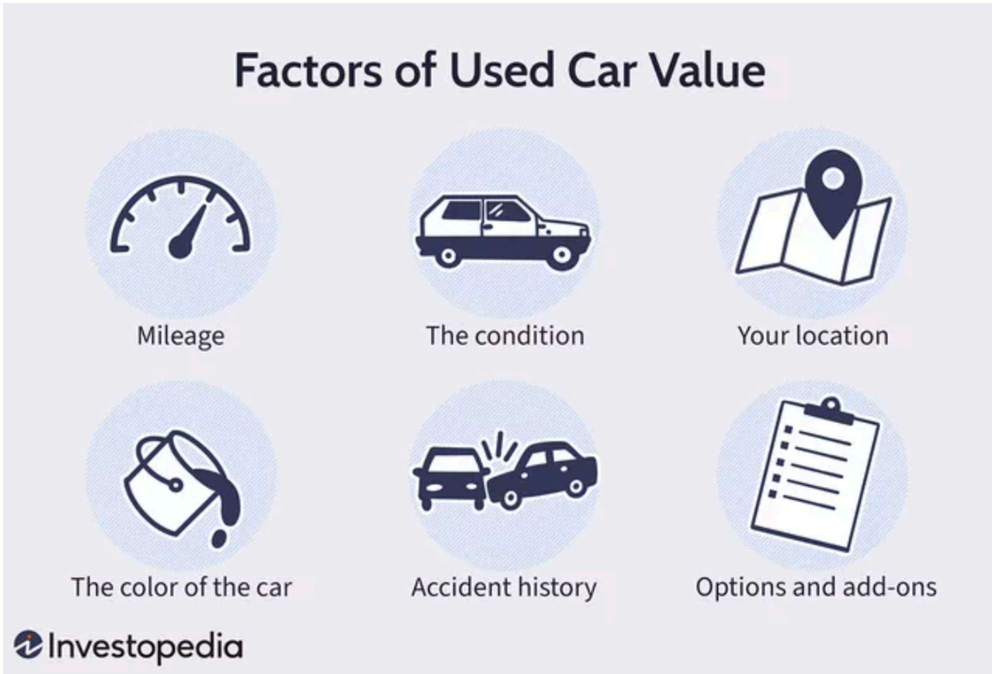

Figure 1: Attributes of the Used Car[1]

either attribute. She has a prior belief about the value of each attribute, and can incur a cost to search for information about the attributes before making a decision. By receiving a noisy signal about an attribute from searching, she can update her belief about the value of that attribute and thus about the value of the product. By assuming that the search cost and the informativeness of the signal are the same for each attribute, we ensure that the attributes are symmetric. So, the consumer will not prefer searching for information about one attribute to the other for exogenous reasons. Which attribute to search at any given time is determined endogenously by the expected gain from an extra piece of information about each attribute.

The consumer will stop searching and buy the good if she becomes optimistic enough about its value (the total value of both attributes), and will stop searching without purchasing if she becomes pessimistic enough about its value. When the consumer's belief about the value of the good is in between, she will search for more information. We characterize the search region by a set of ordinary differential equations for intermediate beliefs and by a system of equations for extreme beliefs. We find that it is always optimal for the consumer to search the attribute about which the consumer has the higher uncertainty due to the faster speed of learning. The consumer only searches for the more uncertain attribute if she

[1] Source of the figure: https://www.investopedia.com/articles/investing/090314/just-what-factors-value-your-used-car.asp

holds a strong prior belief about one of the attributes and may search for both attributes otherwise. In the car purchasing example, a consumer may not bother to search for the safety features of a Volvo car because Volvo has a good reputation for safety. So, she may instead focus on other aspects of the car. In contrast, Faraday Future has not produced any cars yet. If a consumer considers pre-ordering a car, she probably has a lot of uncertainty about everything. So, she may search for information about every attribute.

We study the comparative statics of the optimal search strategy. An increase in the price shifts the entire search region upwards because the consumer needs to gain a higher value from the good to compensate for the higher price. An increase in either the search cost or the noise of the signal makes searching less attractive for the consumer and shrinks the search region.

We also investigate how the consumer's purchasing likelihood depends on the prior belief. When the consumer is optimistic enough about both attributes, she will purchase the product for sure. When she is pessimistic enough about both attributes, she will never purchase the product. When her belief is in between, she will purchase the product with some probability.

In reality, firms can intervene the consumer search and purchase processes by changing consumers' prior beliefs through marketing activities such as advertising. We study the firm's optimal pre-search intervention by assuming that it can disclose the value of one attribute by informative advertising. We find that the firm will not advertise if the consumer's prior beliefs about both attributes are extreme. If the consumer is very optimistic about both attributes, she will purchase the product for sure or with a very high likelihood. So, the firm does not have an incentive to advertise. If the consumer is very pessimistic about both attributes, she will never purchase the product even if she knows that one attribute is good. So, the firm does not advertise either. If the consumer's prior belief is milder, the firm can increase the purchasing probability by advertising. The firm will advertise the better attribute if the consumer is optimistic enough about the worse attribute, and will advertise the worse attribute if the consumer is less optimistic about it.

This paper makes two main contributions. First, we endogenize the search order of different attributes of a product based on the consumer's optimal Bayesian learning. Second, we connect informative advertising with consumer search by comparing the one-dimensional search with advertising and two-dimensional search without advertising.

1.1 Related Literature

This paper is related to the literature on how consumers with limited attention allocate their attention to different attributes or options. Existing literature mainly looks at the case in which the attributes or options are asymmetric (Arbatskaya 2007, Armstrong et al. 2009, Xu et al. 2010, Armstrong and Zhou 2011, Bordolo et al. 2013, Kőszegi and Szeidl 2013, Branco et al. 2016, Zhu and Dukes 2017, Jeziorski and Moorthy 2018). In those papers, consumers know that they face attributes with different prominence/importance ex-ante. For example, the search order is exogenous in Arbatskaya (2007). Armstrong et al. (2009) extend the symmetric search model of Wolinsky (1986) by assuming that there is a prominent firm for which all the consumers will search first. In their model, the prominent firm is exogenous. They do not model why consumers want to search for that firm first. In Bordolo et al. (2013), the salient attribute of a good is the attribute furthest away from the average value of the same attribute in the choice set. In Zhu and Dukes (2017), each competing firm can promote one or both attributes of a product. Though the prominence of the product is endogenously determined by competition, it is exogenously given from the consumer's perspective. Jeziorski and Moorthy (2018) examine the effect of prominence in search advertising. There are two types of prominence in their setting, the position of the ads and the prominence of the advertiser. They find that the ad position prominence and the advertiser prominence are substitutes in consumers' clicking behavior. One of the main contributions of our paper is to endogenize the optimal attribute to search from the consumer's perspective. Instead of assuming that the consumer knows the value of each attribute or learns it at once, as is common in this literature, the Bayesian decision-maker in our model gradually learns the value from noisy signals. So, the relative importance of the attributes may change as the consumer gathers more information. In contrast, the prominence attribute/option in the existing literature does not change over time because they impose exogenous differences on the attributes.

This paper also fits into the literature on optimal information acquisition, particularly consumer search. Stigler 1961 and Weitzman 1979 are among the first papers to derive the optimal search rules under simultaneous and sequential search, respectively. In both papers, the relative importance of different alternatives is exogenous. Consumers observe the distribution of the rewards before making the search decision. Later papers incorporate gradual learning (Moscarini and Smith 2001, Branco et al. 2012, Ke et al. 2016). Like our paper, the attributes are symmetric in those papers. However, the consumer randomly searches for an attribute in those papers. In our model, the consumer decides when to search and which attribute to search. Ke and Villas-Boas (2019) are closely related to our paper. They study the gradual learning of information about multiple alternatives. The decision-maker endogenously determines which alternative to search. There are two main differences between their paper and this one. First, the expected payoff of choosing one of the alternatives depends only on the information gathered from that alternative. So, the objective of searching is to differentiate different alternatives. In our paper, the expected payoff of adopting the product jointly depends on the information gathered from all the alternatives. So, the objective of searching is to learn about the overall distribution of all the attributes. Second, they focus on the decision maker's optimal search strategy. In contrast, we also study the firm's response. We show how the firm can change the consumer's search behavior and increase its profits by pre-search interventions, given the optimal search strategy of the consumer.

Lastly, we study how firms can change consumers' search behavior and increase profits by pre-search interventions such as advertising. People have begun to consider the informational role of advertising since Nelson (1974). Subsequent papers study the disclosure of price (Anderson and Renault 2006) and quality (Lewis and Sappington 1994, Anderson and Renault 2009) by informative advertising. Sun (2011) is the closest paper that studies a seller's disclosure incentive for a product with multiple attributes. It shows that the unraveling result by Grossman (1981) and Milgrom (1981) will not hold if the product has a vertical attribute and a horizontal one. If the product has a high vertical quality, the seller may not disclose the product's horizontal attribute.

Consumers' only source of information about the product comes from the firm in most of the existing advertising literature. In reality, consumers can search for more information after they see the ads. We take it into account by building a micro-founded consumer search model. After the firm advertises, the consumer can still search for information about any attributes. The firm anticipates it when choosing the advertising strategy. Mayzlin and Shin (2011) consider a setting where the consumer can obtain an exogenously given signal by searching for information about the product quality after the firm advertises. Our paper differs from their paper in two ways. On one hand, the quality is vertical in their paper, and the advertising strategy is driven by signaling the firm's private information. We focus on horizontal quality, and the advertising strategy is driven by the difference in one-dimensional search with advertising and two-dimensional search without advertising. On the other hand, the consumer can only search once and observe an aggregate signal about the firm's quality in Mayzlin and Shin (2011). We model the search process in detail so that the consumer chooses what attribute and how long to search. This allows us to endogenize the search order and understand more about the consumer's search behavior and the firm's best response to it.

The remainder of the paper is organized as follows. Section 2 presents the main model. Section 3 studies the comparative statics of the search region. Section 4 characterizes the purchasing likelihood given a prior belief and the consumer's optimal search strategy. Section 5 discusses firms' advertising strategy. Section 6 concludes.

2 Model

A consumer considers whether to purchase a product or not. The product has two attributes whose values are independent. The product's value for the consumer is the sum of the values of the attributes, $U = U_1 + U_2$. The value of each attribute is one if it is good and zero if it is bad. The consumer's prior belief that attribute i is good is $\mu_i(0)$. We assume that the firm does not have private information about the value of the attribute.[2] The price p is exogenously given. We assume that the marginal cost of producing the product is high enough, and thus the price is high enough ($p \geq 3/2$) such that the consumer will quit without purchasing the product for any $\vec{\mu} = (\mu_1, \mu_2)$, if $\mu_1 + \mu_2 \leq 1$. Hence, we restrict our attention to the case in which $\mu_1 + \mu_2 > 1$. The consumer can learn more about the attributes via costly learning before making a decision. At time t, the consumer can make a purchasing decision or search for information. Because of limited attention, she can only search for information about one attribute at a time. So, if the consumer chooses to search for information, she also needs to decide which attribute to search for information about. The game ends when the consumer makes a decision. If the consumer decides to search for information, she will obtain noisy signals about an attribute by incurring a flow cost of c. Define $T_i(t)$ as the cumulative time that attribute i has been searched until time t. We model the signal, S_i, by a Brownian motion (W_i are independent Wiener processes):

$$dS_i(t) = U_i dT_i(t) + \sigma dW_i(T_i(t))$$

The consumer will be more likely to observe a larger signal realization if the attribute is good. Given the received signal, the consumer continuously updates her belief on the value of each attribute according to Bayes' rule.[3] The belief evolution can be characterized by the following ODE:

$$d\mu_i(t) = \frac{1}{\sigma^2}\mu_i(t)[1 - \mu_i(t)]\{dS_i(t) - \mathbb{E}[U_i|\mathcal{F}_t]dT_i(t)\} \tag{1}$$

, where $\{\mathcal{F}_t\}_{t=0}^{+\infty}$ is a filtration with all the observed information up to time t.

The consumer's expected payoff for a given belief $\vec{\mu}$, learning rule α, and stopping time τ is:

$$J(\vec{\mu}, \alpha, \tau) = \mathbb{E}\left\{\max\left[\mu_1(\tau) + \mu_2(\tau) - p, 0\right] - \tau c | \vec{\mu}(0) = \vec{\mu}\right\}$$

[2] This will be more realistic if we consider the horizontal preference rather than the vertical preference.

[3] Notice that the consumer's belief about an attribute will remain the same when she searches for information about the other attribute.

The value function of the consumer's problem is:

$$V(\vec{\mu}) := \sup_{\alpha,\tau} J(\vec{\mu}, \alpha, \tau)$$

Since the learning rule and stopping time should not depend on any future information, the decision at time t should only be based on the observed information up to time $t, \mathcal{F}_t$. It is well known that the current belief $\vec{\mu} = (\mu_1, \mu_2)$ is a sufficient statistic for $\mathcal{F}_t$. So, the learning rule and stopping time will depend only on $\vec{\mu}$. If a learning rule α^* and a stopping time τ^* achieve that value for any given belief, they will be the optimal learning rule and the optimal stopping time.

$$V(\vec{\mu}) = J(\vec{\mu}, \alpha^*, \tau^*)$$

The next section characterizes the consumer's value function and optimal search strategy, including the optimal learning rule and the optimal stopping time.

2.1 Optimal Strategy

When the consumer searches for information about attribute one, the value function satisfies (ignoring the time index t for simplicity):

$$V(\mu_1, \mu_2) = -cdt + \mathbb{E}[V(\mu_1 + d\mu_1, \mu_2)]$$

By Taylor's expansion and Ito's lemma, we get:

$$\frac{\mu_1^2(1-\mu_1)^2}{2\sigma^2} V_{\mu_1\mu_1}(\mu_1, \mu_2) - c = 0 \tag{2}$$

Similarly, when the consumer searches for information about attribute two, we have:

$$\frac{\mu_2^2(1-\mu_2)^2}{2\sigma^2} V_{\mu_2\mu_2}(\mu_1, \mu_2) - c = 0 \tag{3}$$

The HJB equation of the entire problem is:

$$\max\left\{\max_{i=1,2}\left[\frac{\mu_i^2(1-\mu_i)^2}{2\sigma^2} V_{\mu_i\mu_i}(\mu_1, \mu_2) - c\right], \max\left[\mu_1 + \mu_2 - p, 0\right] - V(\mu_1, \mu_2)\right\} = 0 \tag{$\star$}$$

A standard method of solving this kind of stochastic control problem is the"guess and verify" approach. We will show that the value function is a viscosity solution of the HJB equation. Then, we will prove that the viscosity solution of the HJB equation is unique.

Therefore, if we can find a viscosity solution, it must be the value function. To do so, we will construct a learning rule and stopping time, and use it to characterize the search region and the expected payoff. Lastly, we will verify that the conjectured strategy generates a viscosity solution of the HJB equation, which implies that the conjectured learning rule and stopping time are optimal. Because of symmetry, we only need to consider the case in which $\mu_1 \geq \mu_2$. Analytically, we can fully characterize the optimal search strategy when the search cost is low. We do not think the result for the low search cost case is a strong restriction, as we are interested in the consumer's search behavior and how the firm can influence it by informative advertising. Naturally, the more interesting case is when the consumer searches more given the low search cost. When the search cost or the price is very high, the consumer searches little and the problem is less interesting and relevant.

Intuitively, the consumer will stop searching, not buy the product if the belief becomes too low, and will purchase the product if the belief becomes high enough. When the belief is in between, she keeps searching for information. We also conjecture that it is optimal for the consumer to search attribute two, conditional on searching, if $\mu_1 + \mu_2 > 1$ and $\mu_1 \geq \mu_2$.[4] The intuition for this learning rule to be optimal is that the consumer prefers to search for the attribute with a higher rate of learning, as the learning costs are identical. From equation (1), one can see that the more uncertain the belief is, the faster the consumer learns about an attribute. Therefore, she always learns the attribute with a belief closer to 1/2.

Figure 2 illustrates the optimal search strategy. The dashed orange line is the quitting boundary, and the solid blue line is the purchasing boundary. The grey arrow represents which attribute the consumer searches for information about, given the current belief. When the overall beliefs of the attributes are low enough, the likelihood of obtaining lots of positive signals and purchasing the good is too low. The consumer stops searching and quits to save the search cost. When the overall beliefs of the attributes are high enough, purchasing the good gives the consumer a higher enough expected surplus. So, she makes the purchase. In other cases, the consumer searches for more information to make a better decision. Denote the intersection of the quitting boundary and the main diagonal by (μ^*, μ^*), the intersection of the purchasing boundary and the main diagonal by (μ^{**}, μ^{**}). Represent the quitting boundary when $\mu_1 \geq \mu_2$ by $\underline{\mu}(\cdot)$, whose domain is $[\mu^*, 1]$ (the other half of the quitting boundary is determined by symmetry). Represent the purchasing boundary when $\mu_1 \geq \mu_2$ by $\bar{\mu}(\cdot)$, whose domain is $[\mu^{**}, 1]$ (the other half of the purchasing boundary is determined by symmetry).

[4] By symmetry, if $\mu_1 + \mu_2 > 1$ and $\mu_1 < \mu_2$, it is optimal for the consumer to search attribute one, conditional on searching. Note that $\mu_1 + \mu_2 > 1$ always holds in the search region. So, this condition can be omitted. We leave it in the text to emphasize that the consumer searches for information about the attribute with more uncertainty.

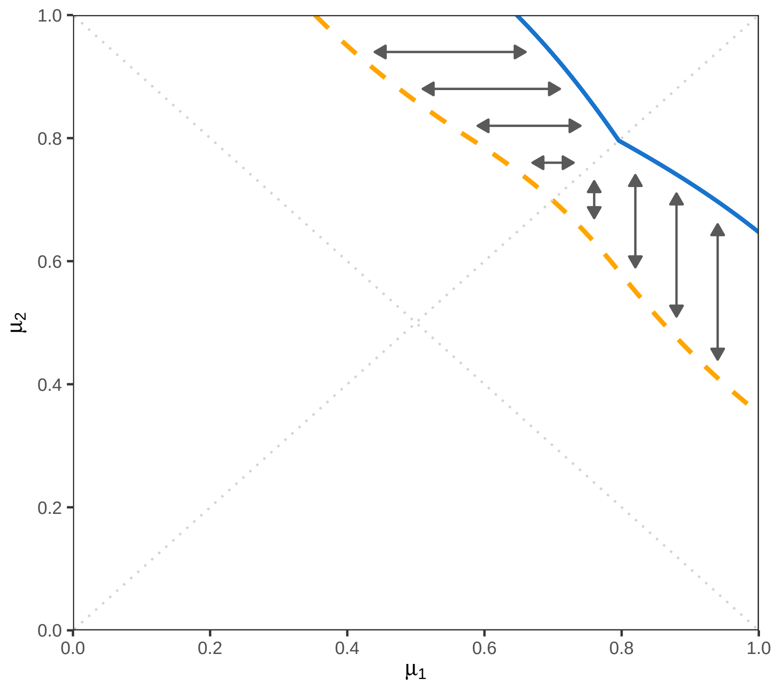

Figure 2: Optimal Search Strategy

The PDE when the consumer searches attribute two, equation (3), has the following general solution:

$$V(\mu_1, \mu_2) = 2\sigma^2 c(1-2\mu_2)\ln\frac{1-\mu_2}{\mu_2} + B_1(\mu_1)\mu_2 + B_2(\mu_1), \mu_1 \in [\mu^*, 1]$$

We also have $V(\mu_1, \mu_2) = 0$ at the quitting boundary $\mu_2 = \underline{\mu}(\mu_1)$. For the value function in the search region, value matching and smooth pasting (wrt μ_2) at the quitting boundary $(\mu_1, \underline{\mu}(\mu_1))$ imply:[5]

$$\frac{V(\mu_1, \mu_2)}{2\sigma^2 c} = (1-2\mu_2)\ln\frac{1-\mu_2}{\mu_2} + \phi(\underline{\mu}(\mu_1))\mu_2 - \psi(\underline{\mu}(\mu_1)) \tag{4}$$

, where $\phi(x) = 2\ln\frac{1-x}{x} + \frac{1}{x} - \frac{1}{1-x}$ and $\psi(x) = \ln\frac{1-x}{x} + \frac{1-2x}{1-x}$.

By symmetry, for $\mu_1 < \mu_2$, the value function in the search region satisfies:

[5] For technical details, please refer to Dixit (1993).

$$\frac{V(\mu_1,\mu_2)}{2\sigma^2 c} = (1-2\mu_1)\ln\frac{1-\mu_1}{\mu_1} + \phi(\underline{\mu}(\mu_2))\mu_1 - \psi(\underline{\mu}(\mu_2)) \tag{5}$$

Equation (4) characterzies the value function for beliefs $\mu_1 \geq \mu_2$. Equation (5) characterzies the value function for beliefs $\mu_1 < \mu_2$. The two regions are separated by the main diagonal $\{(\mu_1,\mu_2) : \mu_1 = \mu_2\}$. Continuity of $V_{\mu_1}(\mu_1,\mu_2)$ at this boundary implies that:

$$\underline{\mu}'(\mu) = \frac{\phi(\underline{\mu}(\mu)) - \phi(\mu)}{\phi'(\underline{\mu}(\mu))[\mu - \underline{\mu}(\mu)]}, \text{ for } \mu \in (\mu^*, \mu^{**}] \tag{D_1}$$

For $\mu_1 \in [\mu^{**}, 1]$, $V(\mu_1,\mu_2) = \mu_1 + \mu_2 - p$ at the purchasing boundary $\mu_2 = \bar{\mu}(\mu_1)$. Value matching and smooth pasting (wrt μ_2) at the purchasing boundary $(\mu_1, \underline{\mu}(\mu_1))$ imply (in the search region):

$$\frac{V(\mu_1,\mu_2)}{2\sigma^2 c} = (1-2\mu_2)\ln\frac{1-\mu_2}{\mu_2} + \phi(\bar{\mu}(\mu_1))\mu_2 - \psi(\bar{\mu}(\mu_1)) + \frac{\mu_1 - \mu_2 - p}{2\sigma^2 c} \tag{6}$$

Equation (4) and (6) use the quitting boundary and the purchasing boundary to pin down the value function, respectively. The resulting expression should be equivalent in the common domain $\mu_1 \in [\mu^{**}, 1]$. By equalizing V and V_{μ_2} of equation (4) and (6), we obtain the following system of equations:

$$\begin{cases} \phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu)) = \frac{1}{2\sigma^2 c} \\ \psi(\underline{\mu}(\mu)) - \psi(\bar{\mu}(\mu)) = \frac{p-\mu}{2\sigma^2 c} \end{cases}, \text{ for } \mu \in [\mu^{**}, 1] \tag{7}$$

For each belief, μ, the system of equations above consists of two unknowns ($\bar{\mu}(\mu)$ and $\underline{\mu}(\mu)$) and two equations. They uniquely determine the function for the purchasing boundary $\bar{\mu}(\mu)$ and the function for the quitting boundary $\underline{\mu}(\mu)$, for $\mu \in [\mu^{**}, 1]$, given a cutoff belief μ^{**}.

Instead of determining $\bar{\mu}(\mu)$ and $\underline{\mu}(\mu)$ by a system of equations (7), we can also implicitly determine $\bar{\mu}(\mu)$ and $\underline{\mu}(\mu)$ in two separate equations. Representing $\bar{\mu}(\mu)$ by $\underline{\mu}(\mu)$ from the first equation of (7), we have:

$$\bar{\mu}(\mu) = \phi^{-1}\left[\phi(\underline{\mu}(\mu)) - \frac{1}{2\sigma^2 c}\right]$$

Plugging it into the second equation of (7), we have:

$$\underline{\mu}(\mu) = \psi^{-1}\left\{\psi\left(\phi^{-1}\left[\phi(\underline{\mu}(\mu)) - \frac{1}{2\sigma^2 c}\right]\right) + \frac{p-\mu}{2\sigma^2 c}\right\}$$

The equation above implicitly determines $\underline{\mu}(\mu)$, for $\mu \in [\mu^{**}, 1]$. Similarly, we can implicitly determine $\bar{\mu}(\mu)$ by the following equation:

$$\bar{\mu}(\mu) = \psi^{-1}\left\{\psi\left(\phi^{-1}\left[\phi(\bar{\mu}(\mu)) + \frac{1}{2\sigma^2 c}\right]\right) - \frac{p-\mu}{2\sigma^2 c}\right\}$$

We now solve for the cutoff belief at the intersection of the purchasing boundary and the main diagonal, μ^{**}. Since (μ^{**}, μ^{**}) is on the purchasing boundary, we have $\mu^{**} = \bar{\mu}(\mu^{**})$, μ^{**} is determined by:

$$\begin{cases} \phi(\underline{\mu}(\mu^{**})) - \phi(\mu^{**}) = \frac{1}{2\sigma^2 c} \\ \psi(\underline{\mu}(\mu^{**})) - \psi(\mu^{**}) = \frac{p-\mu^{**}}{2\sigma^2 c} \end{cases} \tag{8}$$

The system of equations above consists of two unknowns (μ^{**} and $\underline{\mu}(\mu^{**})$) and two equations. They uniquely determine the cutoff belief μ^{**} via the following equations:

$$\phi^{-1}\left[\phi(\mu^{**}) + \frac{1}{2\sigma^2 c}\right] = \psi^{-1}\left[\psi(\mu^{**}) + \frac{p-\mu^{**}}{2\sigma^2 c}\right] \tag{I^{**}}$$

We have pinned down the cutoff belief μ^{**}. Given this cutoff beliefs, we have determined the purchasing boundary $(\mu, \bar{\mu}(\mu))$ and the quitting boundary $(\mu, \underline{\mu}(\mu))$, for $\mu \in [\mu^{**}, 1]$.

The ODE (D_1) and the initial condition (I^{**}) implicitly determine the function for the quitting boundary $\underline{\mu}(\mu)$, for $\mu \in (\mu^*, \mu^{**}]$, given a cutoff belief μ^*.

We now solve for the cutoff belief at the intersection of the quitting boundary and the main diagonal, μ^*. Since (μ^*, μ^*) is on the quitting boundary, we have $\mu^* = \underline{\mu}(\mu^*)$. This initial condition determines μ^*.

In sum, we have pinned down the cutoff belief μ^* and the quitting boundary $(\mu, \underline{\mu}(\mu))$, for $\mu \in [\mu^*, \mu^{**}]$.

We have fully characterized the purchasing boundary $(\mu, \bar{\mu}(\mu))$ and the quitting boundary $(\mu, \underline{\mu}(\mu))$, for $\mu_1 \geq \mu_2$. The other case in which $\mu_1 < \mu_2$ is readily determined by symmetry. The following proposition characterizes the slope of the purchasing and the quitting boundary and the shape of the search region.

Proposition 1. *For $\mu \in (\mu^*, \mu^{**}]$, we have:*

$$\underline{\mu}'(\mu) = \frac{\phi(\underline{\mu}(\mu)) - \phi(\mu)}{\phi'(\underline{\mu}(\mu))[\mu - \underline{\mu}(\mu)]} \tag{D_1}$$

For $\mu \in [\mu^{**}, 1]$, we have:

$$\bar{\mu}'(\mu) = \frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\phi'(\bar{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} \tag{$\overline{D_2}$}$$

$$\underline{\mu}'(\mu) = \frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} \tag{$\underline{D_2}$}$$

Both $\underline{\mu}(\mu)$ and $\bar{\mu}(\mu)$ strictly decrease in μ, while the width of the search region, $\bar{\mu}(\mu) - \underline{\mu}(\mu)$, strictly increases in μ. In addition, if $\underline{\mu}(\mu) \geq 1/2$, then the slope of the quitting boundary is less than -1 and the slope of the purchasing boundary is greater than -1.

We find that the optimal search region has a butterfly shape - the consumer searches for information in a broader region when the consumer is more certain that the more favorable attribute is good. The intuition is the following. The product has a higher expected value if the consumer is more confident about one attribute being good. So, the consumer will search for information about the other attribute even if she has more uncertainty about it. Because the speed of learning is higher when searching a more uncertainty attribute, the benefit of search increases while the search cost remains the same. Therefore, the consumer will search more.

If the consumer likes an attribute more, she will purchase the product even if she has a higher uncertainty about the other attribute. She will also be less likely to stop searching and quit. Therefore, the search region shifts downwards as the belief about one attribute, μ, increases. The value of the slope of the search region is also interesting. It is the marginal rate of substitution between the values of attribute one and two. If the slope equals -1, then the two attributes are perfect substitutes. One may expect this to be the case in general because the product's value is the sum of the values of two attributes. However, both the slope of the quitting boundary and the slope of the purchasing boundary are not -1 in general because of the asymmetry of learning. If the quitting boundary is above 1/2, a unit increase of the belief about attribute one can substitute for more than a unit of the belief about attribute two near the quitting boundary, $\underline{\mu}'(\mu^*) < -1$. The consumer will keep searching for information about attribute two instead of quitting even if μ_2 decreases by slightly more than a unit. This is because the consumer has more uncertainty about attribute 2. The speed of learning is higher when the consumer searches a more uncertainty attribute. So, the benefit of search increases while the search cost remains the same. Similarly, a unit increase of the belief about attribute one can substitute for less than a unit of the belief about attribute two near the purchasing boundary, $\bar{\mu}'(\mu^*) > -1$. The consumer will keep searching for information about attribute two instead of purchasing the product even if μ_2 only decreases by slightly less than a unit.

Given the value function and the optimal strategy derived under the conjectured search strategy, we now verify that the conjectured search strategy is indeed optimal (satisfying the HJB equation ($\star$)).

Theorem 1. *Suppose the search cost is low, $c \leq \frac{1}{2\sigma^2[\phi(1/2)-\phi(\frac{2}{3}p-\frac{1}{6})]}$. Conditional on searching, it is optimal for the consumer to search for information about attribute two (one) if $\mu_1 \geq \mu_2$ ($\mu_1 < \mu_2$).*

We have characterized the search region by a set of ordinary differential equations for moderate beliefs and by a system of equations for extreme beliefs. The optimal search strategy implies that the decision-maker only searches the more uncertain attribute if she holds a strong prior belief on one of the attributes and may search both attributes otherwise. This result is the main testable implication of the paper. Future empirical studies on multi-attribute consumer search can test whether this prediction holds, especially with the aid of the eye-tracking data.

3 Comparative Statics

If the firm wants to use the above results, it needs to understand how the model primitives affect the consumer's search behavior. The following proposition summarizes the comparative statics of the search region with regard to the price, search cost, and noise of the signal.

Proposition 2. *Suppose $\mu_1 \geq \mu_2$. The purchasing threshold $\bar{\mu}(\mu)$ increases in the price p, and decreases in the search cost c and the noise of the signal σ^2. The quitting threshold $\underline{\mu}(\mu)$ increases in the price p, the search cost c, and the noise of the signal σ^2.*

An increase in the price shifts the entire search region upwards because the consumer needs to gain a higher value from the good to compensate for the higher price. For example, as Figure 3 illustrates, the consumer may be willing to pay 1.5 for a good when she believes that each attribute has an 80% probability of being good. She will obtain a positive expected surplus from purchasing the product. However, if the price of the good increases to 1.75, she will not buy the good given the same belief because of the negative expected utility. She may not even keep searching for information because the likelihood that the belief becomes high enough to compensate for the high price is low. She will be better off stopping searching, saving the search cost. Similarly, the consumer may be willing to search for more information when she believes that each attribute has a 70% probability of being good if the price is 1.5. Though she will obtain a negative utility from purchasing the product right away, she may like the product more after some search and gain a positive surplus by purchasing it. In contrast, if the price of the good increases to 1.75, she will stop searching because the likelihood of receiving a lot of positive information and raising the valuation for the product above the high price is very low.

Given a prior belief (μ_1, μ_2), increasing the price has two opposite effects on the firm. A higher price raises the profit conditional on purchasing but reduces the purchasing likelihood. The next section discusses in detail how the consumer's purchasing likelihood depends on the prior belief.

The change in the search cost or the signal's noise has the same effect on the consumer's search behavior because they always appear together in the value function as $c\sigma^2$. An increase in either the search cost or the signal noise makes searching less attractive for the consumer and shrinks the search region. The consumer will only search for information in a narrower range of beliefs. Figure 4 illustrates how the seach region depends on the search cost and the signal noise. For example, for a product whose price is 1.5, the consumer may want to keep searching if she believes that each attribute has a 78% probability of being good and $c\sigma^2 = 0.1$. She can obtain a positive surplus by purchasing the good immediately. However, she may receive some negative information about the product and aviod purchasing a bad product by mistake. So, she may prefer to make a deicsion when she becomes more certain about the value of the product. However, if it takes more time or effort to search for information or the information is not very accurate, $c\sigma^2 = 0.2$, the benefit from search will be lower and the consumer may instead purchase the good immediately.

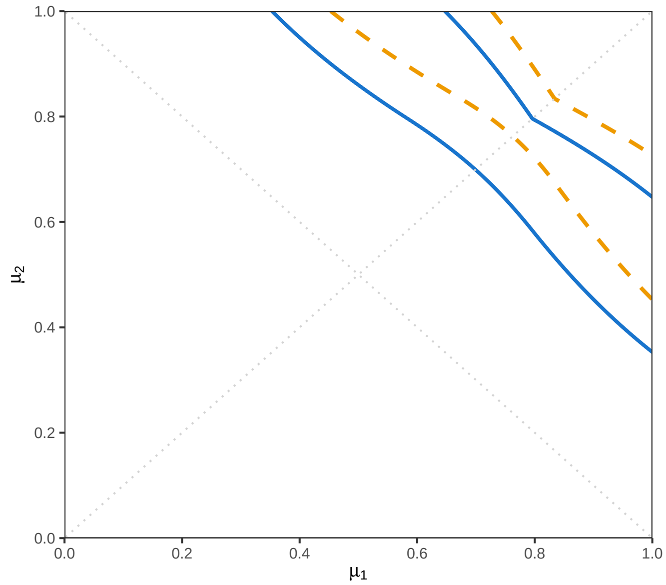

Figure 3: Optimal Search Region for $p = 1.5$ (solid blue) or 1.6 (dashed orange), $c = 0.1, \sigma^2 = 1$.

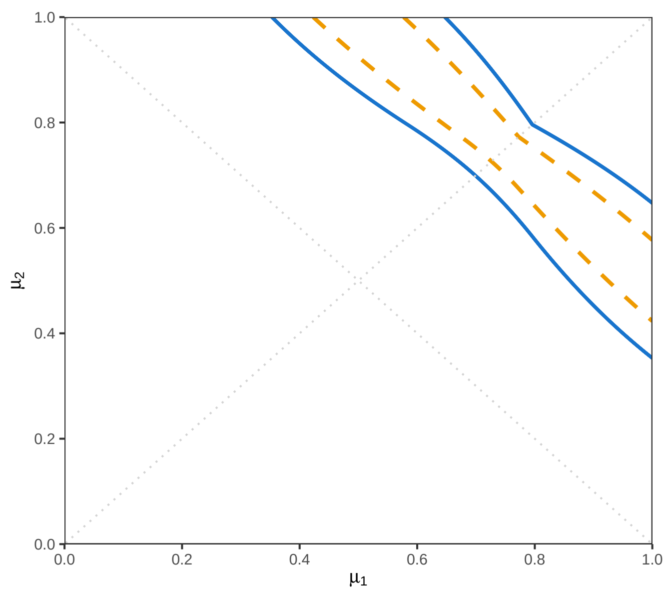

Figure 4: Optimal Search Region for $p = 1.5$, $c\sigma^2 = 0.1$ (solid blue) or 0.2 (dashed orange).

4 Purchasing Likelihood

We now look at the consumer's belief path to purchase. If the consumer strongly believes that one of the attributes is good, she will never search for information on that attribute. The consumer will keep searching for information about the other attribute. She will purchase the product if she obtains enough positive information and the belief reaches the purchasing boundary $\bar{\mu}$. If she receives enough negative information and the belief reaches the quitting boundary $\underline{\mu}$, she will quit searching without buying the good. For example, when deciding whether to buy a Volvo, a consumer may not bother to search for its safety features because Volvo has a good reputation for safety. She gains more from searching for other attributes of the car.

In contrast, the consumer must search for information on both attributes before purchasing the good if she has mild beliefs about both attributes. Moreover, she will be equally certain about the value of each attribute if she decides to buy the good. For example, Faraday Future has not produced any cars yet. If a consumer considers pre-ordering a car, she probably has a lot of uncertainty about everything. So, she may search for information about every attribute. Given the consumer's optimal search strategy, we can calculate the purchasing likelihood given a prior belief (μ_1, μ_2).

Proposition 3. *Suppose $\mu_1 \geq \mu_2$. The probability that the consumer purchases the product is:*

$$\begin{aligned} P(\mu_1, \mu_2) &:= \mathbb{P}[\textit{purchasing}|\textit{starting at } (\mu_1, \mu_2)] \\ &= \begin{cases} 1, \ \textit{if } \mu_1 \in [\mu^{**}, 1] \textit{ and } \mu_2 \in [\bar{\mu}(\mu_1), \mu_1] \\ \frac{\mu_2 - \underline{\mu}(\mu_1)}{\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)}, \ \textit{if } \mu_1 \in [\mu^{**}, 1] \textit{ and } \mu_2 \in [\underline{\mu}(\mu_1), \bar{\mu}(\mu_1)] \\ h(\mu_1, \mu_2)\tilde{P}(\mu_1), \ \textit{if } \mu_1 \in [\mu^{*}, \mu^{**}] \textit{ and } \mu_2 \in [\underline{\mu}(\mu_1), \mu_1] \\ 0, \ \textit{if } \mu_1 \leq \mu^{*} \textit{ or } \mu_2 \leq \bar{\mu}(\mu_1) \end{cases} \end{aligned}$$

, where $h(\mu_1, \mu_2) = \frac{\mu_2 - \underline{\mu}(\mu_1)}{\mu_1 - \underline{\mu}(\mu_1)}$ and $\tilde{P}(\mu) = e^{-\int_{\mu}^{\mu^{**}} \frac{2}{x - \underline{\mu}(x)} dx}$. By symmetry, $P(\mu_1, \mu_2) = P(\mu_2, \mu_1)$ if $\mu_1 < \mu_2$.

We can see that there are four regions, as Figure 5 illustrates. The consumer makes the purchase immediately if the belief lies in the region $S1$ and quits without purchasing immediately if the belief lies in the region $S4$. For beliefs in between, the value of information is the highest. The consumer will search for more information before making a decision. If the belief lies in the region $S3$ on the right-hand side of the figure, the consumer strongly believes that the first attribute is good. So, instead of spending more time confirming it, she searches for information about the more uncertain attribute, attribute two. If she receives enough positive information about the second attribute, she will be very optimistic about the product's value and will make the purchase. If she receives enough negative information about the second attribute, she will be pessimistic about the product's value and will stop searching. Because the conumer has had a pretty good sense of the first atribute's value, she will not switch back to searching for information about it regardless of what she learns about the second attribute. Therefore, the second attribute is the pivotal attribute in this case.

If the belief lies on the right-hand side of the region $S2$, the consumer is quite uncertain about the value of both attributes. She will search for information about attribute two because she is more uncertain about attribute two than attribute one. However, the consumer also does not have a strong belief about the value of attribute one. So, the consumer will switch to search for information about attribute one if she receives enough positive signals about attribute two. She may switch back to attribute two if she gets enough positive signals about attribute one and may switch back and forth before being confident about both attributes and purchasing the product. As shown in Figure 5, the belief must reach (μ^{**}, μ^{**}) for the consumer to make the purchase decision. So, she will be equally confident about the value of both attributes when she stops searching and buying the good. She will stop searching and quit if she receives enough negative signals about either attribute.

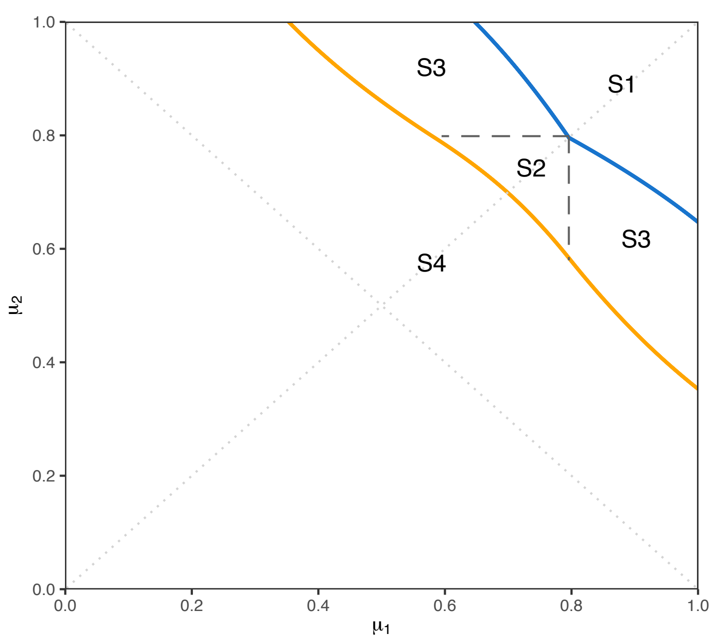

Figure 5: Four Regions for Purchase

5 Pre-search Interventions

The previous section determines the purchasing probability given the prior belief. In reality, firms can intervene the consumer search and purchase processes through marketing activities such as advertising. The firm can reveal the value of the attribute by *informative advertising*. The consumer does not need to incur costs to search for information about the attribute. Due to the limited bandwidth of ads, we assume that the firm can only reveal the value of one attribute.

5.1 Informative Advertising

By conducting informative advertising, the firm can disclose the value of one of the attributes. Given the updated information, the consumer can search for more information before making a decision. If the firm advertises one attribute, it reveals its value. So, the consumer only has uncertainty about the other attribute. Her search problem becomes a single-attribute problem. Suppose the firm advertises attribute $i \in \{1, 2\}$. The value of attribute i, U_i, becomes 1 with probability μ_i and 0 with probability $1 - \mu_i$.[6] The consumer can make a decision right away or search for information about attribute $j := 3 - i$. The real price of the product is $p' := p - U_i$. One can see that the consumer will quit if $U_i = 0$. So, we consider the case in which $U_i = 1$ now (p' becomes $p - 1$). The optimal search strategy has been shown in Branco et al. (2012) and Ke and Villas-Boas (2019). There exists $0 < \underline{\mu}_j < \bar{\mu}_j < 1$ such that the consumer searches for more information if $\mu_j \in (\underline{\mu}_j, \bar{\mu}_j)$, purchases the product if $\mu_j \geq \bar{\mu}_j$, and quits if $\mu_j \leq \underline{\mu}_j$. In the search region, the value function is determined by:

$$
\begin{aligned}
&\frac{\mu_j^2(1-\mu_j)^2}{2\sigma^2} W''(\mu_j) - c = 0 \\
\Rightarrow\ & W(\mu_j) = 2\sigma^2 c(1-2\mu_j)\ln\frac{1-\mu_j}{\mu_j} + K_1\mu_j + K_2,\ \ \mu_j \in (\underline{\mu}_j, \bar{\mu}_j)
\end{aligned}
$$

Since $W(\underline{\mu}_j) = W'(\underline{\mu}_j) = 0$, $W(\bar{\mu}_j) = \bar{\mu}_j - p'$, and $W'(\bar{\mu}_j) = 1$, value matching and smooth pasting at $\underline{\mu}_j$ and $\bar{\mu}_j$ determine the cutoff belief:

$$
\begin{cases}
\phi(\underline{\mu}_j) - \phi(\bar{\mu}_j) = \frac{1}{2\sigma^2 c} \\
\psi(\underline{\mu}_j) - \psi(\bar{\mu}_j) = \frac{p-1}{2\sigma^2 c}
\end{cases}
\tag{9}
$$

[6] We denote $\mu_i(0)$ by μ_i to simplify the notation in this section.

By symmetry, we only need to consider the firm's advertising strategy when $\mu_1 \geq \mu_2$, which is summarized by the following proposition.

Proposition 4. *Suppose $\mu_1 \geq \mu_2$. There exists $\tilde{\mu}(\mu_1)$ and $\widehat{\mu}(\mu_1)$ such that $\underline{\mu}(1) < \tilde{\mu}(\mu_1) \leq \widehat{\mu}(\mu_1) < \bar{\mu}(\mu_1)$ and $\tilde{\mu}(\mu_1)$ decreases in μ_1. The firm does not advertise if $\mu_1 \leq \underline{\mu}(1)$ or $\mu_2 \geq \widehat{\mu}(\mu_1)$, advertises attribute two if $\mu_1 \in (\underline{\mu}(1), \bar{\mu}(1)]$, or $\mu_1 > \bar{\mu}(1)$ and $\mu_2 \leq \tilde{\mu}(\mu_1)$, advertises attribute one if $\mu_1 > \bar{\mu}(1)$ and $\mu_2 \in (\tilde{\mu}(\mu_1), \widehat{\mu}(\mu_1))$.*

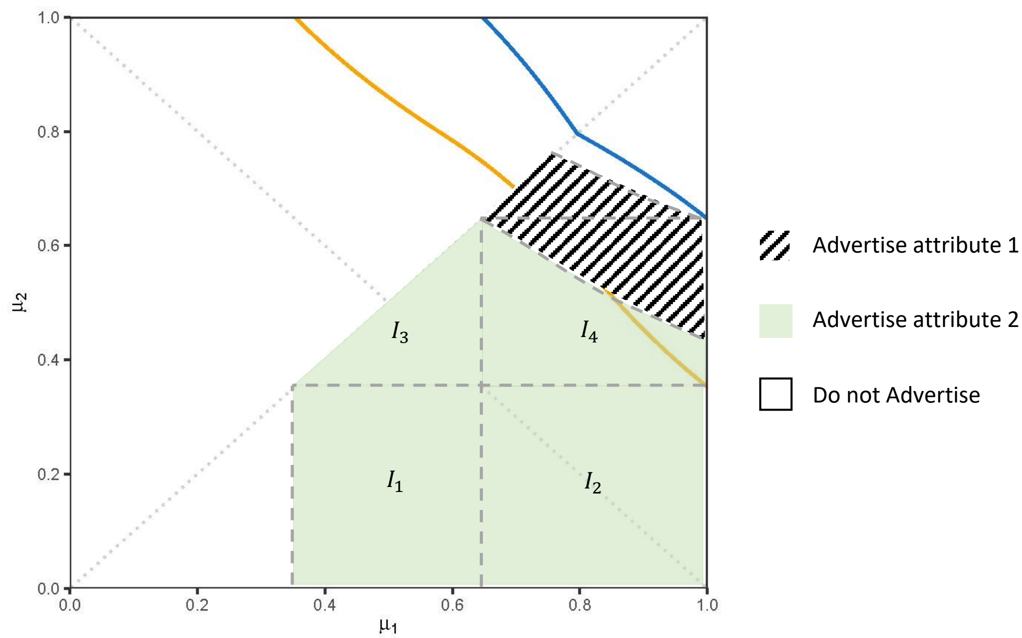

Figure 6: Advertising Strategy

Figure 6 illustrates the advertising strategy. The firm advertises attribute one in the diagonal striped black region, attribute two in the solid green region, and does not advertise in the white region. If the consumer's prior beliefs about both attributes are too low, the product will not be attractive to the consumer even if she knows that one attribute is good. The consumer will neither search for information nor purchase the product even if the firm advertises. So, the firm does not advertise. If the consumer has high enough prior beliefs about both attributes, she will purchase the product without searching. The firm also has no incentive to advertise. Even if the consumer's belief is within the search region, she will purchase the product after receiving a little positive information as long as her belief is close to the purchasing boundary. The purchasing probability is close to 1. In contrast, if the firm advertises, the consumer will quit for sure if she finds out that one attribute is bad. So, the purchasing probability is lower. The firm is better off by not advertising. The intuition is the following. If the consumer finds out that one attribute is good from advertising, her belief about the product value will be higher than what is needed for her to purchase the product immediately. Such excessive belief is wasteful from the firm's standpoint. If the firm does not advertise, the consumer will be just indifferent between searching for more information and purchasing the product after receiving a little positive information. The firm does not waste any belief. Therefore, the consumer will be more likely to purchase the product without advertising. Therefore, the firm does not advertise in the white region.

Now let's consider the solid green region and the diagonal striped black region. We divide the solid green region into four sub-regions. If the belief lies in the region I_1 or I_2, the consumer is very pessimistic about the second attribute. Even if she knows for sure that the first attribute is good, she needs to receive a lot of positive signals about attribute two to purchase the product. The search cost outweighs the benefit of the search. So, she will not search for information. The only way of inducing the consumer to search is to advertise attribute two. With a high probability, the consumer will find out that attribute two is bad and quit. However, if the consumer find out that attribute two is good, she needs fewer positive signals to purchase the product by searching for attribute one. The benefit of search outweighs the search cost. So, the consumer will search for information about attribute one and purchase the product with a positive probability. Therefore, the firm advertises attribute two.

If the belief lies in the region I_3, the consumer will never purchase the product without advertising but may purchase the product if the firm advertises either attribute. So, the firm advertises. Since the consumer is more optimistic about attribute one, she will be more likely to search for information if the firm advertises attribute one than two. However, she needs more positive signals to purchase the product. So, the conversion rate conditional on searching is lower. It turns out that the second effect is stronger than the first effect. So, the firm advertises attribute two.

Lastly, we consider the case where the belief lies in the region I_4 or the diagonal striped black region. If the consumer's belief about attribute two is high enough, she will purchase the product immediately if she knows attribute one is good. One can see that the firm always prefers to advertise attribute one to attribute two. If the consumer's belief about attribute two is lower, the comparison between advertising attribute one and two is non-trivial. If the firm advertises attribute one and the consumer finds out it is good, the consumer will always search for information about attribute two before making a decision. In contrast, the consumer will be very positive about the product value if the firm advertises attribute two and the consumer knows that attribute two is good. In that case, she will purchase the product immediately. So, some beliefs are "wasted" - the consumer will purchase the product immediately even if her belief is lower. The more optimistic she is about the first attribute, the more beliefs are wasted. So, the firm will be more likely to advertise attribute one.

In sum, the firm will not advertise if the consumer's prior beliefs about both attributes are extreme and will advertise if the consumer's prior belief is milder. In that case, the firm will advertise the better attribute if the consumer is optimistic enough about the worse attribute, and will advertise the worse attribute if the consumer is less optimistic about it.

5.2 Advertising Costs

In the previous discussion, we did not consider the advertising costs. In reality, the firm needs to incur a cost to advertise. Our framework can incorporate this cost, but the analysis will be more tedious. So, we abstract away the advertising costs in the previous analysis. We briefly discuss what happens if we take into account the advertising costs. Suppose the firm needs to incur a cost c_A to advertise attribute i. The comparison between advertising attribute one and two will not change because both require an extra cost, c_A. However, whether the firm prefers to advertise or not may change. If the prior belief of the consumer without advertising is close to the purchasing boundary, then the firm will not advertise. Even without advertising, the consumer will purchase the product with a high probability. By not advertising, the firm saves advertising costs. The firm will also not advertise if the belief about one of the attributes is too low. Even if the firm can raise the purchasing probability above zero by advertising, the purchasing likelihood is very low. The profit will be negative because of the advertising costs. So, the firm will not advertise, and the consumer will neither search nor purchase. For all other beliefs, the firm's advertising strategy is the same as the case without advertising costs.

6 Conclusion

Understanding how consumers decide which attribute to pay more attention to has important managerial implications. It helps the firm decide how to design the product and which attributes to emphasize. In this paper, we study the optimal search strategy of a Bayesian decision-maker by endogenizing the optimal attribute to search for, when to keep searching, and when to stop and make a decision. We characterize the search region by a set of ordinary differential equations for moderate beliefs and by a system of equations for extreme beliefs. We find that it is always optimal to search the attribute the consumer has the highest uncertainty due to the fastest learning speed. The decision-maker only searches the more uncertain attribute if she holds a strong prior belief on one of the attributes, and may search both attributes otherwise. We also study the firm's optimal pre-search intervention by assuming that it can disclose the value of one attribute by informative advertising. We find that the firm will not advertise if the consumer's prior beliefs about both attributes are extreme. If the consumer is very optimistic about both attributes, she will purchase the product for sure or with a very high likelihood. So, the firm does not have an incentive to advertise. If the consumer is very pessimistic about both attributes, she will never purchase the product even if she knows that one attribute is good. So, the firm does not advertise either. If the consumer's prior belief is milder, the firm can increase the purchasing probability by advertising. The firm will advertise the better attribute if the consumer is optimistic enough about the worse attribute, and will advertise the worse attribute if the consumer is less optimistic about it.

There are some limitations to this paper. The consumer only considers one product in our model. If there are multiple products, the consumer needs to make two decisions - which product to search for and which attribute of the product to search for. Studying this richer problem can lead to interesting findings. It will also be interesting to extend the number of attributes beyond two and see whether the consumer still searches for the attribute with the highest uncertainty due to the fastest learning speed. Lastly, we consider an exogenous price throughout the paper to focus on the role of information. Future research can study the optimal pricing of the product given the consumer's optimal search strategy.

Appendix

Proof of Proposition 1. We have derived (D_1) in the main text. It implies immediately that $\underline{\mu}'(\mu) < 0$ for $\mu \in (\mu^*, \mu^{**}]$. For $\mu \in [\mu^{**}, 1]$, by the implicit function theorem, we have:

$$
\begin{aligned}
\begin{bmatrix} \bar{\mu}'(\mu) \\ \underline{\mu}'(\mu) \end{bmatrix} &= -\begin{bmatrix} -\phi'(\bar{\mu}(\mu)) & \phi'(\underline{\mu}(\mu)) \\ -\psi'(\bar{\mu}(\mu)) & \psi'(\underline{\mu}(\mu)) \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 \\ \frac{1}{2\sigma^2 c} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2\sigma^2 c} \frac{1}{\phi'(\bar{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]} \\ \frac{1}{2\sigma^2 c} \frac{1}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\phi(\underline{\mu}(\mu))-\phi(\bar{\mu}(\mu))}{\phi'(\bar{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]} < 0 \\ \frac{\phi(\underline{\mu}(\mu))-\phi(\bar{\mu}(\mu))}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]} < 0 \end{bmatrix}
\end{aligned}
$$

This gives us the expression for ($\overline{D_2}$) and ($\underline{D_2}$). One can see from the negative sign of the derivative that both $\underline{\mu}(\mu)$ and $\bar{\mu}(\mu)$ strictly decrease in μ.

We now look at the width of the search region.

$$
\begin{aligned}
&[\bar{\mu}(\mu) - \underline{\mu}(\mu)]' \\
=&\frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\phi'(\bar{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} - \frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} \\
=&\frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\bar{\mu}(\mu) - \underline{\mu}(\mu)} \left[1/\phi'(\bar{\mu}(\mu)) - 1/\phi'(\underline{\mu}(\mu))\right] \\
=&\frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\bar{\mu}(\mu) - \underline{\mu}(\mu)} \left[\underline{\mu}(\mu)^2(1-\underline{\mu}(\mu))^2 - \bar{\mu}(\mu)^2(1-\bar{\mu}(\mu))^2\right]
\end{aligned}
$$

One can see that $\frac{\phi(\underline{\mu}(\mu))-\phi(\bar{\mu}(\mu))}{\bar{\mu}(\mu)-\underline{\mu}(\mu)} > 0$. So, $[\bar{\mu}(\mu) - \underline{\mu}(\mu)]' > 0 \Leftrightarrow \underline{\mu}(\mu)^2(1-\underline{\mu}(\mu))^2 > \bar{\mu}(\mu)^2(1-\bar{\mu}(\mu))^2 \Leftrightarrow \underline{\mu}(\mu)(1-\underline{\mu}(\mu)) > \bar{\mu}(\mu)(1-\bar{\mu}(\mu)) \Leftrightarrow |\underline{\mu}(\mu) - 1/2| < |\bar{\mu}(\mu) - 1/2|$. Thus, the width of the search region, $\bar{\mu}(\mu) - \underline{\mu}(\mu)$, increases in the belief, μ, if and only if the quitting boundary is closer to 1/2 than the purchasing boundary. We know that $\forall \mu \geq \mu^{**},\ p = \mu + \frac{\bar{\mu}(\mu)+\underline{\mu}(\mu)}{2}$ due to the symmetry of the one-dimensional learning problem.[7] Therefore,

[7] More specifically, the sum of the purchasing and quitting thresholds is zero when the price is zero in the one-dimensional optimal search strategy, as shown by Branco et al. (2012). It implies that the price equals to the average of the two boundaries. In our two-dimensional problem, the consumer only searches the more uncertain attribute when $\mu \geq \mu^{**}$. So, it can be translated to a one-dimensional search problem with the price p normalized to $p - \mu$.

$$
\begin{aligned}
&\frac{\bar{\mu}(\mu)+\underline{\mu}(\mu)}{2} = p-\mu \geq 3/2-1 = 1/2 \\
\Rightarrow\ &\frac{\bar{\mu}(\mu)+\underline{\mu}(\mu)}{2} \geq 1/2 \\
\Leftrightarrow\ &\bar{\mu}(\mu)+\underline{\mu}(\mu) > 1 \\
\Leftrightarrow\ &|\underline{\mu}(\mu)-1/2| < |\bar{\mu}(\mu)-1/2|,\ \forall \mu \geq \mu^{**}
\end{aligned}
$$

Thus, the width of search region, $\bar{\mu}(\mu)-\underline{\mu}(\mu)$, always increases in the belief μ.

Now suppose that $\underline{\mu}(\mu) \geq 1/2$, then $\forall \mu \in (\mu^*, \mu^{**}]$, we have

$$
\begin{aligned}
\underline{\mu}'(\mu) &\overset{(D_1)}{=} \frac{\phi(\underline{\mu}(\mu))-\phi(\mu)}{\phi'(\underline{\mu}(\mu))[\mu-\underline{\mu}(\mu)]} \\
&= \frac{-\phi'(\xi_1(\mu))[\mu-\underline{\mu}(\mu)]}{\phi'(\underline{\mu}(\mu))[\mu-\underline{\mu}(\mu)]}\ (\xi_1(\mu) \in (\underline{\mu}(\mu), \mu)) \\
&= -\frac{\phi'(\xi_1(\mu)}{\phi'(\underline{\mu}(\mu))} \\
&< -1
\end{aligned}
$$

, where the last inequality comes from the fact that the absolute value of $\phi'(x) = -\frac{1}{x^2(1-x)^2}$ is strictly increasing in x for $x \geq 1/2$. Similarly, $\forall \mu \in [\mu^{**}, 1]$, we have

$$
\begin{aligned}
\underline{\mu}'(\mu) &\overset{(\underline{D_2})}{=} \frac{\phi(\underline{\mu}(\mu))-\phi(\bar{\mu}(\mu))}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]} \\
&= \frac{-\phi'(\xi_2(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]}\ (\xi_2(\mu) \in (\underline{\mu}(\mu), \bar{\mu}(\mu))) \\
&= -\frac{\phi'(\xi_2(\mu)}{\phi'(\underline{\mu}(\mu))} \\
&< -1 \\
\bar{\mu}'(\mu) &\overset{(\overline{D_2})}{=} \frac{\phi(\underline{\mu}(\mu))-\phi(\bar{\mu}(\mu))}{\phi'(\bar{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]} \\
&= \frac{-\phi'(\xi_3(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]}{\phi'(\bar{\mu}(\mu))[\bar{\mu}(\mu)-\underline{\mu}(\mu)]}\ (\xi_3(\mu) \in (\underline{\mu}(\mu), \bar{\mu}(\mu))) \\
&= -\frac{\phi'(\xi_3(\mu))}{\phi'(\bar{\mu}(\mu))} \\
&> -1
\end{aligned}
$$

□

Proof of Theorem 1. By symmetry, we only need to prove the case of $\mu_1 \geq \mu_2$. We first show that the viscosity solution of the HJB equation ($\star$) exists and is unique. Since the value function is a viscosity solution of ($\star$), the viscosity solution of ($\star$) must be the value function by uniqueness. We then just need to verify that the learning strategy we conjectured indeed generates a viscosity solution to ($\star$). So, the conjectured strategy is optimal.

Lemma 1. *The viscosity solution of the HJB equation ($\star$) exists and is unique.*

Proof. Since the consumer can guarantee a payoff of zero by quitting immediately and cannot achieve a payoff higher than $\sup\{\mu_1 + \mu_2 - p\} = 1 + 1 - p \leq 2$, the value function is bounded and thus exists. This implies the existence of the viscosity solution because the value function is a viscosity solution to ($\star$).

The proof of the uniqueness uses a modification of a comparison principle in Crandall et al. (1992). Given that it very much resembles the proof of Lemma 1 in Ke and Villas-Boas (2019), we refer the reader to their proof. □

To verify that the conjectured strategy indeed generates a viscosity solution to the HJB equation ($\star$):

$$\max\left\{\max_{i=1,2}\left[\frac{\mu_i^2(1-\mu_i)^2}{2\sigma^2}V_{\mu_i\mu_i}(\mu_1,\mu_2) - c\right], \max\left[\mu_1+\mu_2-p, 0\right] - V(\mu_1,\mu_2)\right\} = 0$$

We just need to show that (everything else holds by our construction):

$$\begin{aligned} &\frac{\mu_1^2(1-\mu_1)^2}{2\sigma^2}V_{\mu_1\mu_1}(\mu_1,\mu_2) - c \leq 0 \\ \Leftrightarrow\ &\mu_1^2(1-\mu_1)^2 V_{\mu_1\mu_1}(\mu_1,\mu_2)/2\sigma^2 c \leq 1 \end{aligned} \tag{10}$$

if $\mu_1 + \mu_2 > 1$, $\mu_1 \geq \mu_2$, and $\underline{\mu}(\mu_1) < \mu_2 < \bar{\mu}(\mu_1)$.

For $\mu_1 \in (\mu^*, \mu^{**}]$, we have

$$
\begin{aligned}
& V_{\mu_1}(\mu_1, \mu_2)/2\sigma^2 c \\
=& \phi'(\underline{\mu}(\mu_1))\underline{\mu}'(\mu_1)[\mu_2 - \underline{\mu}(\mu_1)] \\
\overset{(D_1)}{=}& \frac{\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)}{\mu_1 - \underline{\mu}(\mu_1)}[\mu_2 - \underline{\mu}(\mu_1)] \\
\Rightarrow\ & V_{\mu_1\mu_1}(\mu_1, \mu_2)/2\sigma^2 c \\
=& \phi'(\underline{\mu}(\mu_1))\underline{\mu}'(\mu_1)[\mu_2 - \underline{\mu}(\mu_1)] \\
\overset{(D_1)}{=}& \frac{\phi'(\underline{\mu}(\mu_1))\underline{\mu}'(\mu_1) - \phi'(\mu_1)}{\mu_1 - \underline{\mu}(\mu_1)}[\mu_2 - \underline{\mu}(\mu_1)] + [\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)]\frac{(\mu_2 - \mu_1)\underline{\mu}'(\mu_1) + \underline{\mu}(\mu_1) - \mu_2}{[\mu_1 - \underline{\mu}(\mu_1)]^2} \\
=& -\frac{\phi'(\mu_1)[\mu_2 - \underline{\mu}(\mu_1)]}{\mu_1 - \underline{\mu}(\mu_1)} + (\mu_2 - \mu_1)\frac{[\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)]^2}{\phi'(\underline{\mu}(\mu_1))[\mu_1 - \underline{\mu}(\mu_1)]^3} \\
\Rightarrow\ & \mu_1^2(1-\mu_1)^2 V_{\mu_1\mu_1}(\mu_1, \mu_2)/2\sigma^2 c \\
=& \frac{\mu_2 - \underline{\mu}(\mu_1)}{\mu_1 - \underline{\mu}(\mu_1)} + (\mu_1 - \mu_2)\mu_1^2(1-\mu_1)^2\frac{\underline{\mu}(\mu_1)^2[1 - \underline{\mu}(\mu_1)]^2}{[\mu_1 - \underline{\mu}(\mu_1)]^3}[\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)]^2
\end{aligned}
$$

So,

$$
\begin{aligned}
& \mu_1^2(1-\mu_1)^2 V_{\mu_1\mu_1}(\mu_1, \mu_2)/2\sigma^2 c \leq 1 \\
\Leftrightarrow & \mu_1^2(1-\mu_1)^2\frac{\underline{\mu}(\mu_1)^2[1 - \underline{\mu}(\mu_1)]^2}{[\mu_1 - \underline{\mu}(\mu_1)]^2}[\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)]^2 \leq 1 \\
\Leftrightarrow & \mu_1(1-\mu_1)\frac{\underline{\mu}(\mu_1)^2[1 - \underline{\mu}(\mu_1)]}{[\mu_1 - \underline{\mu}(\mu_1)]}[\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)] \leq 1 \\
\Leftrightarrow & H(\mu_1) := \mu_1(1-\mu_1)[\phi(\underline{\mu}(\mu_1)) - \phi(\mu_1)] - \frac{\mu_1 - \underline{\mu}(\mu_1)}{\underline{\mu}(\mu_1)[1 - \underline{\mu}(\mu_1)]} \leq 0
\end{aligned} \tag{11}
$$

Observe that $H(\mu^*) = 0$. Ignoring the subscript 1 for notational ease, we have:

$$
\begin{aligned}
H'(\mu) =& (1-2\mu)[\phi(\underline{\mu}(\mu)) - \phi(\mu)] + \frac{\mu(1-\mu)}{\mu - \underline{\mu}(\mu)}[\phi(\underline{\mu}(\mu)) - \phi(\mu)] + \frac{1}{\mu(1-\mu)} \\
& - \frac{1}{\underline{\mu}(\mu)(1 - \underline{\mu}(\mu))} + \frac{\phi(\underline{\mu}(\mu)) - \phi(\mu)}{\mu - \underline{\mu}(\mu)}[-\mu + 2\mu\underline{\mu}(\mu) - \underline{\mu}(\mu)^2] \\
=& [1 - 3\mu + \underline{\mu}(\mu)][\phi(\underline{\mu}(\mu)) - \phi(\mu)] + \frac{1}{\mu(1-\mu)} - \frac{1}{\underline{\mu}(\mu)(1 - \underline{\mu}(\mu))}
\end{aligned}
$$

Suppose (11) does not hold. There would exist $\widehat{\mu}$ such that $H(\widehat{\mu}) = 0$ and $H'(\widehat{\mu}) > 0$.

$$(11) \Rightarrow \phi(\underline{\mu}(\widehat{\mu})) - \phi(\widehat{\mu}) = \frac{\widehat{\mu} - \underline{\mu}(\widehat{\mu})}{\widehat{\mu}(1-\widehat{\mu})\underline{\mu}(\widehat{\mu})[1-\underline{\mu}(\widehat{\mu})]}$$

Hence, we get an expression for $\frac{1}{\widehat{\mu}(1-\widehat{\mu})}$ and $\frac{1}{\underline{\mu}(\mu)[1-\underline{\mu}(\mu)]}$. Plugging these expressions into the previous expression for $H'(\mu)$, we have:

$$H'(\widehat{\mu}) = -2[\phi(\underline{\mu}(\mu)) - \phi(\mu)][\mu - \underline{\mu}(\mu)] \leq 0$$

A contradiction! So, (11) and thus (10) hold, $\forall \mu_1 \in [\mu^*, \mu^{**}]$.

For $\mu_1 \in [\mu^{**}, 1]$, we have

$$\begin{aligned}
V_{\mu_1}(\mu_1,\mu_2)/2\sigma^2 c =& \phi'(\underline{\mu}(\mu_1))\underline{\mu}'(\mu_1)[\mu_2 - \underline{\mu}(\mu_1)] \\
\overset{(\underline{D_2})}{\equiv} & \frac{\mu_2 - \underline{\mu}(\mu_1)}{\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)} \\
V_{\mu_1\mu_1}(\mu_1,\mu_2)/2\sigma^2 c =& \frac{-\underline{\mu}'(\mu_1)[\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)] - [\bar{\mu}'(\mu_1) - \underline{\mu}'(\mu_1)][\mu_2 - \underline{\mu}(\mu_1)]}{[\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)]^2} \\
=& \frac{1}{2\sigma^2 c}\frac{1}{[\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)]^3}[\frac{\mu_2 - \bar{\mu}(\mu_1)}{\phi'(\underline{\mu}(\mu_1))} - \frac{\mu_2 - \underline{\mu}(\mu_1)}{\phi'(\bar{\mu}(\mu_1))}] \\
\Rightarrow V_{\mu_1\mu_1}(\mu_1,\mu_2) =& \frac{1}{[\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)]^3}[\frac{\mu_2 - \bar{\mu}(\mu_1)}{\phi'(\underline{\mu}(\mu_1))} - \frac{\mu_2 - \underline{\mu}(\mu_1)}{\phi'(\bar{\mu}(\mu_1))}]
\end{aligned}$$

Since $\frac{\partial V_{\mu_1\mu_1}(\mu_1,\mu_2)}{\partial \mu_2} < 0$, we only need to show that (10) holds for $\mu_2 = \underline{\mu}(\mu_1)$:

$$\begin{aligned}
&\mu_1^2(1-\mu_1)^2 V_{\mu_1\mu_1}(\mu_1,\underline{\mu}(\mu_1))/2\sigma^2 c \leq 1 \\
\Leftrightarrow & \frac{\mu_1^2(1-\mu_1)^2}{[\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)]^2}\frac{-1}{\phi'(\underline{\mu}(\mu_1))} \leq 1
\end{aligned} \tag{12}$$

Let's first show that $\underline{\mu}(\mu^{**}) \leq 1/2$ by contradiction. Suppose instead $\underline{\mu}(\mu^{**}) > 1/2$.

$$\begin{aligned}
p - \mu^{**} &= \frac{\bar{\mu}(\mu^{**}) + \underline{\mu}(\mu^{**})}{2} \\
\Leftrightarrow\ p - \mu^{**} &= \frac{\mu^{**} + \underline{\mu}(\mu^{**})}{2} \\
\Leftrightarrow\ \underline{\mu}(\mu^{**}) &= 2p - 3\mu^{**}
\end{aligned}$$

Hence, $2p - 3\mu^{**} > 1/2 \Rightarrow \mu^{**} < \frac{2}{3}p - \frac{1}{6}$. Since $\phi(x)$ is strictly decreasing in x, the first equation of (8) implies

$$
\begin{aligned}
\frac{1}{2\sigma^2 c} =& \phi(\underline{\mu}(\mu^{**})) - \phi(\mu^{**}) \\
<& \phi(1/2) - \phi(\frac{2}{3}p - \frac{1}{6}) \\
\Leftrightarrow c >& \frac{1}{2\sigma^2[\phi(1/2) - \phi(\frac{2}{3}p - \frac{1}{6})]}
\end{aligned}
$$

A contradiction! Therefore, $\underline{\mu}(\mu^{**}) \leq 1/2$. Since $\underline{\mu}(\mu_1)$ is decreasing in μ_1, we have $\underline{\mu}(\mu_1) \leq 1/2$, $\forall \mu \in [\mu^{**}, 1]$. One can see that the LHS of (12), $\frac{\mu_1^2(1-\mu_1)^2}{[\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)]^2} \frac{-1}{\phi'(\underline{\mu}(\mu_1))}$, decreases in $\mu_1 \in [\mu^{**}, 1]$. And we know that (12) holds for $\mu_1 = \mu^{**}$ (we have shown that (10) and thus (12) hold for $\forall \mu_1 \in [\mu^*, \mu^{**}]$). Therefore, (12) and thus (10) hold for $\forall \mu_1 \in [\mu^{**}, 1]$.

□

Proof of Proposition 2.

(1) Comparative statics w.r.t. p

We first consider $\bar{\mu}(\mu)$. Fixing an arbitrary $\mu \in (\mu^{**}, 1]$, recall the system of equations (7):

$$
\begin{aligned}
\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu)) &= \frac{1}{2\sigma^2 c} \\
\psi(\underline{\mu}(\mu)) - \psi(\bar{\mu}(\mu)) &= \frac{p - \mu}{2\sigma^2 c}
\end{aligned}
$$

By the implicit function theorem, we obtain:

$$
\begin{aligned}
\begin{bmatrix} \frac{\partial \bar{\mu}(\mu)}{\partial p} \\ \frac{\partial \underline{\mu}(\mu)}{\partial p} \end{bmatrix} &= -\begin{bmatrix} -\phi'(\bar{\mu}(\mu)) & \phi'(\underline{\mu}(\mu)) \\ -\psi'(\bar{\mu}(\mu)) & \psi'(\underline{\mu}(\mu)) \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 \\ -\frac{1}{2\sigma^2 c} \end{bmatrix} \\
&= \begin{bmatrix} -\frac{\phi(\bar{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} & > 0 \\ -\frac{\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu))}{\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} & > 0 \end{bmatrix}
\end{aligned}
$$

We now consider $\underline{\mu}(\mu)$. Suppose there exists $p_1 > p_2$ with the corresponding quitting boundaries $(\mu_1, \underline{\mu}_{p_1}(\mu_1))$ and $(\mu_1, \underline{\mu}_{p_2}(\mu_1))$, respectively. Denote the cutoff beliefs by $(\mu^*_{p_1}, \mu^{**}_{p_1})$ for price p_1 and by $(\mu^*_{p_2}, \mu^{**}_{p_2})$ for price p_2. Fixing an arbitrary $\mu_1 \in (\mu^*_{p_1}, 1]$, we know that the consumer is indifferent between quitting and searching for information when her belief is $(\mu_1, \underline{\mu}_{p_1}(\mu_1))$ and the price is p_1. Since $p_2 < p_1$, one can see that the value of searching for information when her belief is $(\mu_1, \underline{\mu}_{p_1}(\mu_1))$ and the price is p_2 is strictly higher than zero. So, the consumer will keep searching for information. Thus, $\underline{\mu}_{p_2}(\mu_1) < \underline{\mu}_{p_1}(\mu_1)$.

Therefore, the entire search region shifts upwards as the price increases.

(2) Comparative statics w.r.t. c

We first consider $\bar{\mu}(\mu)$. Fixing an arbitrary $\mu \in (\mu^{**}, 1]$, recall the system of equations (7):

$$\phi(\underline{\mu}(\mu)) - \phi(\bar{\mu}(\mu)) = \frac{1}{2\sigma^2 c}$$
$$\psi(\underline{\mu}(\mu)) - \psi(\bar{\mu}(\mu)) = \frac{p-\mu}{2\sigma^2 c}$$

By the implicit function theorem, we obtain:

$$\begin{bmatrix} \frac{\partial \bar{\mu}(\mu)}{\partial c} \\ \frac{\partial \underline{\mu}(\mu)}{\partial c} \end{bmatrix} = - \begin{bmatrix} -\phi'(\bar{\mu}(\mu)) & \phi'(\underline{\mu}(\mu)) \\ -\psi'(\bar{\mu}(\mu)) & \psi'(\underline{\mu}(\mu)) \end{bmatrix}^{-1} \cdot \begin{bmatrix} \frac{1}{2\sigma^2 c^2} \\ \frac{p-\mu}{2\sigma^2 c^2} \end{bmatrix}$$
$$= \frac{1}{2\sigma^2 c^2 \phi'(\bar{\mu}(\mu))\phi'(\underline{\mu}(\mu))[\bar{\mu}(\mu) - \underline{\mu}(\mu)]} \cdot \begin{bmatrix} \phi'(\underline{\mu}(\mu))(p - \mu - \underline{\mu}(\mu)) \\ \phi'(\bar{\mu}(\mu))(p - \mu - \bar{\mu}(\mu)) \end{bmatrix}$$

The consumer purchases the product when the belief is $(\mu, \bar{\mu}(\mu))$. So, $\mu + \bar{\mu}(\mu) - p > 0$. The consumer stops searching and does not purchase the product when the belief is $(\mu, \underline{\mu}(\mu))$. So, $\mu + \underline{\mu}(\mu) - p < 0$. We also have $\phi'(x) = -\frac{1}{x^2(1-x)^2} \Rightarrow \phi'(x) < 0, \forall x$. Thus, we obtain:

$$\begin{bmatrix} \frac{\partial \bar{\mu}(\mu)}{\partial c} < 0 \\ \frac{\partial \underline{\mu}(\mu)}{\partial c} > 0 \end{bmatrix}$$

We now consider $\underline{\mu}(\mu)$. Suppose there exists $c_1 > c_2$ with the corresponding quitting boundaries $(\mu_1, \underline{\mu}_{c_1}(\mu_1))$ and $(\mu_1, \underline{\mu}_{c_2}(\mu_1))$, respectively. Denote the cutoff beliefs by $(\mu^*_{c_1}, \mu^{**}_{c_1})$ for price c_1 and by $(\mu^*_{c_2}, \mu^{**}_{c_2})$ for price c_2. Fixing an arbitrary $\mu_1 \in (\mu^*_{c_1}, 1]$, we know that the consumer is indifferent between quitting and searching for information when her belief is $(\mu_1, \underline{\mu}_{c_1}(\mu_1))$ and the price is c_1. Since $c_2 < c_1$, one can see that the value of searching for information when her belief is $(\mu_1, \underline{\mu}_{c_1}(\mu_1))$ and the price is c_2 is strictly higher than zero. So, the consumer will keep searching for information. Thus, $\underline{\mu}_{c_2}(\mu_1) < \underline{\mu}_{c_1}(\mu_1)$.

(3) Comparative statics w.r.t. σ^2

c and σ^2 always appear together as $2\sigma^2 c$ in the equations. So, the qualitative result of the comparative statics w.r.t. σ^2 is the same as the comparative statics w.r.t. c.

$\square$

Proof of Proposition 3. We first consider $\mu_1 \in [\mu^{**}, 1]$ and $\mu_1 \geq \mu_2$. Under this circumstance, the consumer only learns about attribute two until μ_2 hits either the purchasing boundary or the quitting boundary. As μ_2 is a martingale, by Dynkin's formula, we get:

$$P(\mu_1, \mu_2) := \mathbb{P}[\text{purchasing}|\text{starting at } (\mu_1, \mu_2)] = \frac{\mu_2 - \underline{\mu}(\mu_1)}{\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)} \tag{13}$$

Now we consider $\mu_1 \in [\mu^*, \mu^{**}]$ and $\mu_1 \geq \mu_2$. The belief either hits (μ^{**}, μ^{**}) and the consumer purchases the good or the belief hits $\{(x, \underline{\mu}(x)) : x \in [\mu_1, \mu^{**})\} \cup \{(\underline{\mu}(x), x) : x \in [\mu_1, \mu^{**})\}$ and the consumer quits. To calculate the purchasing likelihood, let's first calculate the likelihood of the belief hitting $(\mu_1, \underline{\mu}(\mu_1))$ before hitting the main diagonal (μ_1, μ_1), $q(\mu_1, \mu_2)$.

$$q(\mu_1, \mu_2) = \frac{\mu_1 - \mu_2}{\mu_1 - \underline{\mu}(\mu_1)}$$

Now we calculate the probability of purchasing given belief (μ, μ), $\tilde{P}(\mu)$ by consider the infinitesimal learning on attribute two. Noticing that $q(\mu, \mu) = 0$, $\frac{\partial q}{\partial \mu_1}|_{\mu_1=\mu_2=\mu} = \frac{1}{\mu - \underline{\mu}(\mu)}$, $\frac{\partial q}{\partial \mu_2}|_{\mu_1=\mu_2=\mu} = -\frac{1}{\mu - \underline{\mu}(\mu)}$, we have:

$$\begin{aligned}
\tilde{P}(\mu) &= \frac{1}{2}\mathbb{P}[\text{purchasing}|(\mu, \mu), d\mu \geq 0] + \frac{1}{2}\mathbb{P}[\text{purchasing}|(\mu, \mu), d\mu < 0] \\
&= \frac{1}{2}[1 - q(\mu + |d\mu|, \mu)]\tilde{P}(\mu + |d\mu|) + \frac{1}{2}[1 - q(\mu - |d\mu|, \mu)]\tilde{P}(\mu) \\
&= \tilde{P}(\mu) + \frac{|d\mu|}{2}\tilde{P}'(\mu) + |d\mu|\frac{\tilde{P}(\mu)}{\underline{\mu}(\mu) - \mu} + o(d\mu) \\
\Rightarrow 0 &= \frac{|d\mu|}{2}\left[\tilde{P}'(\mu) + 2\frac{\tilde{P}(\mu)}{\underline{\mu}(\mu) - \mu}\right] + o(d\mu) \\
\Rightarrow \frac{\tilde{P}'(\mu)}{\tilde{P}(\mu)} &= -\frac{2}{\underline{\mu}(\mu) - \mu}, \ \forall \mu \in (\mu^*, \mu^{**})
\end{aligned}$$

, where the last equality comes from dividing the previous equation by $|d\mu|$ and take the limit of $d\mu$ to 0. Together with the initial condition $\tilde{P}(\mu^{**}) = 1$, we obtain:

$$\tilde{P}(\mu) = e^{-\int_{\mu}^{\mu^{**}} \frac{2}{x - \underline{\mu}(x)} dx}$$

In sum, the purchasing likelihood when $\mu_1 \geq \mu_2$ and $\mu_1 \in (\mu^*, \mu^{**})$ is:

$$P(\mu_1, \mu_2) = \mathbb{P}[\text{purchasing}|\text{starting at } (\mu_1, \mu_2)] = [1 - q(\mu_1, \mu_2)]\tilde{P}(\mu_1) = h(\mu_1, \mu_2)\tilde{P}(\mu_1) \tag{14}$$

, where $h(\mu_1, \mu_2) = \frac{\mu_2 - \underline{\mu}(\mu_1)}{\mu_1 - \underline{\mu}(\mu_1)}$.

By symmetry, the purchasing likelihood when $\mu_1 < \mu_2$ and $\mu_2 \in (\mu^*, \mu^{**})$ is:

$$P(\mu_1, \mu_2) = P(\mu_2, \mu_1) = [1 - q(\mu_2, \mu_1)]\tilde{P}(\mu_2) = h(\mu_2, \mu_1)\tilde{P}(\mu_2) \tag{15}$$

□

Proof of Proposition 4. One can see that the consumer will not purchase the product if $\mu_1 \leq \underline{\mu}(1)$, even if the firm advertises one attribute which turns out to be good. So, the firm does not advertise if $\mu_1 \leq \underline{\mu}(1)$. Also, the consumer will purchase the product for sure if $\mu_2 \geq \bar{\mu}(\mu_1)$ without advertising. So, the firm does not advertise if $\mu_2 \geq \bar{\mu}(\mu_1)$. We now look at other cases.

(1) $\mu_1 > \underline{\mu}(1)$ and $\mu_2 \leq \underline{\mu}(1)$ (Region I_1 and I_2)

The consumer will never purchase the product if the firm advertises attribute one or does not advertise. In contrast, the consumer may purchase theproduct if the firm advertises on attribute two. The consumer will not purchase if attribute two is bad. However, if attribute two is good, the consumer will purchase the product immediately in the region I_2, and will search for information about attribute one in the region I_1. In the region I_1, the consumer will purchase the product after receiving enough positive information. So, the purchasing likelihood is strictly positive. Hence, the firm advertises attribute two.

(2) $\mu_1 \in (\underline{\mu}(1), \bar{\mu}(1)]$ and $\mu_2 > \underline{\mu}(1)$ (Region I_3)

The purchasing probability is zero if the firm does not advertise, and is positive if the firm advertises either attriutes. Thus, we need to compare the purchasing likelihoods between advertising attribute one and two. We use $P_i(\mu_1, \mu_2)$ to denote the purchasing probability when the prior belief is (μ_1, μ_2) and the firm advertises attribute i.

$$
\begin{aligned}
P_1(\mu_1,\mu_2) =&\ \mu_1 \cdot \frac{\mu_2 - \underline{\mu}(1)}{\bar{\mu}(1) - \underline{\mu}(1)} \\
P_2(\mu_1,\mu_2) =&\ \mu_2 \cdot \frac{\mu_1 - \underline{\mu}(1)}{\bar{\mu}(1) - \underline{\mu}(1)} \\
&\overset{\mu_1 \geq \mu_2}{\geq} P_1(\mu_1,\mu_2)
\end{aligned}
$$

, where the inequality is strict if $\mu_1 > \mu_2$. So, the firm advertises attribute two.

(3) $\mu_1 > \bar{\mu}(1)$ and $\mu_2 \in (\underline{\mu}(1), \bar{\mu}(\mu))$ (Region I_4, the diagonal striped black region, and the white search region)

To characterize the advertising strategy, we need to determine two things. First, whether the firm wants to advertise. Second, whether the firm prefers advertising attribute one or two, conditional on advertising.

We first compare advertising attribute one and two.

$$
\begin{aligned}
P_1(\mu_1,\mu_2) =&\ \mu_1 \cdot \frac{\mu_2 - \underline{\mu}(1)}{\bar{\mu}(1) - \underline{\mu}(1)} \\
P_2(\mu_1,\mu_2) =&\ \mu_2 \\
P_1(\mu_1,\mu_2) > P_1(\mu_1,\mu_2) \Leftrightarrow&\ \frac{\mu_2 - \underline{\mu}(1)}{\bar{\mu}(1) - \underline{\mu}(1)} > \frac{\mu_2}{\mu_1} \\
\Leftrightarrow&\ \mu_2 > \frac{\underline{\mu}(1)\mu_1}{\mu_1 - \bar{\mu}(1) + \underline{\mu}(1)} := \tilde{\mu}(\mu_1)
\end{aligned}
$$

So, the firm prefers advertising attribute one to advertising attribute two if and only if $\mu_2 > \tilde{\mu}(\mu_1)$. One can see that $\tilde{\mu}(\mu_1)$ decreases in μ_1.

We then determine whether the firm wants to advertise or not. If the belief is below the purchasing boundary, the firm always prefers advertising because the consumer will never purchase without advertising. Now suppose the belief is in the search region, $\mu_1 \in [\mu^{**}, 1]$ and $\mu_2 \in [\bar{\mu}(\mu_1), \mu_1]$. According to Proposition 3, the purchasing likelihood without advertising is:

$$
P(\mu_1,\mu_2) = \frac{\mu_2 - \underline{\mu}(\mu_1)}{\bar{\mu}(\mu_1) - \underline{\mu}(\mu_1)}
$$

If the firm advertises attribute one, the purchasing likelihood is:

$$P_1(\mu_1, \mu_2) = \begin{cases} \mu_1 \cdot \frac{\mu_2 - \underline{\mu}(1)}{\bar{\mu}(1) - \underline{\mu}(1)}, & \text{if } \mu_2 < \bar{\mu}(1) \\ \mu_1, & \text{if } \mu_2 \geq \bar{\mu}(1) \end{cases}$$

If the firm advertises attribute two, the purchasing likelihood is:

$$P_2(\mu_1, \mu_2) = \mu_2$$

Observe that $P(\mu_1, \underline{\mu}(\mu)) = 0$, $P(\mu_1, \underline{\mu}(\mu)) = 1$, $P_1(\mu_1, \underline{\mu}(1)) = 0$, $P_1(\mu_1, \bar{\mu}(\mu)) = \mu_1$, and $\underline{\mu}(1) \leq \underline{\mu}(\mu)$. By (quasi-) linearity of the purchasing likelihood, one can see that $P(\mu_1, \mu_2)$ crosses $P_1(\mu_1, \mu_2) \vee P_2(\mu_1, \mu_2)$ once as μ_2 increases, fixing a μ_1. Hence, there exists $\widehat{\mu}(\mu_1) \in [\tilde{\mu}(\mu_1), \bar{\mu}(\mu_1))$ such that the firm does not advertise if and only if $\mu_2 \geq \widehat{\mu}(\mu_1)$.

$\square$

References

Anderson, S. P., Renault, R. (2006). Advertising content. American Economic Review, 96(1), 93-113.

Anderson, S. P., Renault, R. (2009). Comparative advertising: disclosing horizontal match information. The RAND Journal of Economics, 40(3), 558-581.

Arbatskaya, M. (2007). Ordered search. The RAND Journal of Economics, 38(1), 119-126.

Armstrong, M., Zhou, J. (2011). Paying for prominence. The Economic Journal, 121(556), F368-F395.

Armstrong, M., Vickers, J., Zhou, J. (2009). Prominence and consumer search. The RAND Journal of Economics, 40(2), 209-233.

Bordalo, P., Gennaioli, N., Shleifer, A. (2013). Salience and consumer choice. Journal of Political Economy, 121(5), 803-843.

Branco, F., Sun, M., Villas-Boas, J. M. (2012). Optimal search for product information. Management Science, 58(11), 2037-2056.

Branco, F., Sun, M., Villas-Boas, J. M. (2016). Too much information? Information provision and search costs. Marketing Science, 35(4), 605-618.

Crandall, M. G., Ishii, H., Lions, P. L. (1992). User's guide to viscosity solutions of second order partial differential equations. Bulletin of the American mathematical society, 27(1), 1-67.

Dixit, A. (1993). The art of smooth pasting. Harwood Academic Publishers: Chur, Switzerland.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. The Journal of Law and Economics, 24(3), 461-483.

Jeziorski, P., Moorthy, S. (2018). Advertiser prominence effects in search advertising. Management science, 64(3), 1365-1383.

Ke, T. T., Villas-Boas, J. M. (2019). Optimal learning before choice. Journal of Economic Theory, 180, 383-437.

Ke, T. T., Shen, Z. J. M., Villas-Boas, J. M. (2016). Search for information on multiple products. Management Science, 62(12), 3576-3603.

Kőszegi, B., Szeidl, A. (2013). A model of focusing in economic choice. The Quarterly journal of economics, 128(1), 53-104.

Lewis, T. R., Sappington, D. E. (1994). Supplying information to facilitate price discrimination. International Economic Review, 309-327.

Mayzlin, D., Shin, J. (2011). Uninformative advertising as an invitation to search. Marketing science, 30(4), 666-685.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. The Bell Journal of Economics, 380-391.

Moscarini, G., Smith, L. (2001). The optimal level of experimentation. Econometrica, 69(6), 1629-1644.

Nelson, P. (1974). Advertising as information. Journal of political economy, 82(4), 729-754.

Stigler, G. J. (1961). The economics of information. Journal of political economy, 69(3), 213-225.

Sun, M. (2011). Disclosing multiple product attributes. Journal of Economics & Management Strategy, 20(1), 195-224.

Weitzman, M. L. (1979). Optimal search for the best alternative. Econometrica: Journal of the Econometric Society, 641-654.

Wolinsky, A. (1986). True monopolistic competition as a result of imperfect information. The Quarterly Journal of Economics, 101(3), 493-511.

Xu, L., Chen, J., Whinston, A. (2010). Oligopolistic pricing with online search. Journal of Management information systems, 27(3), 111-142.

Zhu, Y., Dukes, A. (2017). Prominent attributes under limited attention. Marketing Science, 36(5), 683-698.